

6

Structural aspects of language endangerment

Naomi Palosaari and Lyle Campbell

6.1 Introduction

This chapter addresses structural aspects of language endangerment from two perspectives: the contributions that the study of endangered languages make to typology and linguistic theory, and the structural consequences of language endangerment, including the kinds of changes that can take place in the phonology, morphology and syntax of endangered languages.

6.2 Contributions from endangered languages to typology and linguistic theory

A major goal of linguistics is to understand what is possible and what is impossible in human languages, and through this to understand the potentials and limitations of human cognition as reflected in language. It is important to obtain scientific information about endangered languages, for if they are lost without documentation, we stand to lose valuable information about the full range of human languages, about their parts and patterns, structures and uses, and how these things interact with one another. To illustrate this point, we present several examples of STRUCTURAL PHENOMENA (from both sound systems and morphosyntax) which have been uncovered from work on endangered languages and which could well have never been known if these languages had disappeared before they were documented and described. We illustrate how such findings contribute to language typology and linguistic theory. TYPOLGY, broadly speaking, is the classification of languages according to linguistic traits and the comparison or classification of linguistic features and structures across languages.

More specifically, typology is understood in different ways; among them:

- the classification of structural types cross-linguistically;
- the investigation of cross-linguistic generalizations concerning patterns among linguistic traits; and
- as a general approach to linguistics which attempts to explain the patterns and classification through appeal to language function in cross-linguistic comparison; that is, the relation between linguistic form and function.

Typology is closely associated with the study of LINGUISTIC UNIVERSALS, which can be understood as the common characteristics of the world's languages, usually with the goal of providing insight into the fundamental nature of human language. Thus discoveries in typology and the identification of universals contribute to linguistic theory, which is aimed at understanding and explaining the nature of human language.¹ Section 6.2.1 discusses sounds and sound systems, and 6.2.2 explores morphosyntactic examples.

6.2.1 Sounds and sound systems

6.2.1.1 Unique speech sounds: a Nivacle case study

Nivacle (Chulupí), an endangered Matacoan language with c. 250 speakers in Argentina and c. 8,500 in Paraguay, has a speech sound not found in any other language.² It is a complex segment composed of a voiceless velar stop and a voiced alveolar lateral resonant, articulated and released simultaneously (that is, with both articulatory gestures formed at the same time and released as nearly simultaneously as possible), represented as $/k̠l/$ (Campbell and Grondona 2007, Campbell in preparation).³ The discovery of a new speech sound is for linguists like the discovery of a new species is for biologists. We document and describe endangered languages precisely so information of this sort will not be lost to science. The discovery of this unique sound has considerable typological significance: we must add a new sound to the inventory of speech sounds in human languages, and its discovery has implications for general claims about languages.

The phonemic inventory of Nivacle consonants is given in Table 6.1.

Nivacle provides counterexamples to a number of proposed cross-linguistic generalizations about laterals and liquids, offering valuable evidence about the possible structure of sound systems. For example, Maddieson (1984: 88) proposed that:

- (1) a language with two or more liquids is expected to have a contrast between a lateral and a non-lateral. However, in Nivacle both liquids are laterals and there are no non-lateral liquids (no 'r' sounds).

Table 6.1. *Nivacle consonant inventory*

plain stops/affricates	p	t	ts	č	k	ʔ
glottalized stops/affricates ⁴	p'	t'	ts'	č'	k'	
fricatives	ɸ		s	š	x	
lateral		l				
velar + lateral		kl				
nasals	m	n				

- (2) a language with one or more laterals typically has one voiced lateral approximant: for example, a marked lateral in a language implies the presence also of plain 'l', and voiceless 'l' also implies the presence of plain 'l'. This, however, is not true of Nivacle; although it has two laterals, one is a voiceless approximant $/l̥/$ ⁵ and the other, $/kl̥/$, is not an approximant. Nivacle has no plain (voiced) $/l/$.
- (3) a language with two or more laterals is expected to contrast them either in point of articulation or in manner of articulation, but not in both. The Nivacle laterals, however, differ both in point of articulation and manner of articulation.

Thus, the Nivacle laterals illustrate well how the discovery of a new speech sound in the investigation of an endangered language can have an impact on general claims about language. Given the counter-examples just mentioned, all these proposed generalizations need to be reevaluated.

6.2.1.2 The Ubykh sound inventory

Ubykh is a now-extinct language of the North-West Caucasian family, formerly spoken in Turkey. Ubykh had perhaps the largest consonant inventory in the world – aside from some of the Khoisan click languages – with 81 documented native consonants, including glottalized, pharyngealized, labialized and palatalized series, and distinguishing at least 8 points of articulation, with some 27 distinct fricatives and 20 uvular consonants (Dumézil 1965, Vogt 1963). Ubykh was reasonably well described before becoming extinct, and so we know of its unusual inventory of consonants and of the fine-grained phonemic contrasts which it shows can exist in a single language. This is information that would be lost to science if it had become extinct without having been described. Still, it is almost certain that much more information of scientific value was irretrievably lost when Ubykh ceased to be spoken.

6.2.1.3 Other unique sounds

Other examples can easily be cited of previously unknown sounds that are now known to linguistic science through recent documentation and description of endangered languages. For example, Ladefoged and Everett (1996) made known the occurrence of a sound in the Chapakuran

languages Wari' and Oro Win (Brazil) composed of a voiceless laminal dental plosive followed by a bilabial trill, [t̪β], an allophone of /t/ before /o/ and /u/. Pirahã (Muran family, Brazil, with about 100 speakers) has two highly unusual sounds, a voiced bilabial trill [β] (rare in other languages) and a lateral-apical double-flap [ɺɺ] (unique to Pirahã). The first is an allophone of /b/, the latter of /g/ (Everett 1984).

Again, such findings have implications for general claims about language. For example, some have claimed there is a connection between the size of a language's consonant inventory and the kinds of consonants expected to be in it. According to Lindblom and Maddieson's (1988) 'size principle', smaller inventories of consonants tend to contain only consonants which are simpler (to produce or to perceive) and that more complex consonants are found in languages with larger consonant inventories.⁶ This is certainly challenged by the examples cited here, since Pirahã has an extremely small phonemic inventory, with only eleven phonemes, and Wari' has only twelve. It is all too plausible that these languages of Amazonia could have become extinct without leaving any traces,⁷ giving us no inkling that such sounds are possible in human languages and resulting in erroneous theorizing about constraints on possible sound systems in human languages.

Other examples of sounds discovered in endangered languages include the linguo-labial segments (tongue tip or blade against the upper lip, which is drawn downward to meet the tongue) in a group of languages in Vanuatu (on Espiritu Santo and Malekula) and in Umotina (extinct, Bororan family; Brazil) (Ladefoged and Maddieson 1996:18–19, 43–4); and voiceless implosives, known now from, for example, Owerri Igbo, Uzere Isoko, and several languages of the K'ichean branch of Mayan (Campbell 1973, Ladefoged and Maddieson 1996: 87–9). Until the discovery of these sounds it was believed that implosives had to be voiced – as Maddieson (1984: 121) put it: 'an imploded segment is voiced' (see also Ladefoged and Maddieson 1996: 87–9).

6.2.1.4 Xinkan vowel harmony

Xinkan is a small family of four languages in Guatemala (not demonstrably related to any other languages or language families). Two of the languages recently became extinct and the other two each have only one or two semi-speakers (people who have learned the language imperfectly and are not fully fluent). Nevertheless, three of the four languages have been investigated rather intensively, resulting in reasonably extensive documentation and description (Rogers *et al.* in preparation).⁸ Xinkan has a unique VOWEL HARMONY pattern (restrictions on which vowels can co-occur with one another in the same word). In most languages with vowel harmony, vowels within a word agree in frontness vs. backness or in roundness vs. non-roundness, or in having advanced tongue root (ATR) or lacking ATR. However, in Xinkan languages, the harmonic sets of vowels depend on vowel height, but with exceptional behaviour for

/i/ (Rogers 2008). The harmonic sets of vowels (in which a member of a set can co-occur in a word with other members of that set, but not with vowels of the other sets) are: {i, u, a}, {e, o, a}, and {i, a}. That is, within roots the high and mid vowels cannot co-occur, the low vowel /a/ is a neutral vowel which can occur with vowels of any height in any of the sets, and the high central vowel /i/ can only co-occur with other instances of /i/ or with neutral /a/.

- (1) {i, u, a}

hiiru	‘monkey’
ts’am’u	‘close your eyes’
pari	‘day’
ts’uuḥi	‘non-Indian person’
- (2) {o, e, a}

k’oosek	‘large’
seema	‘fish’
goona	‘hill’
- (3) {i, a}

ts’iim’aḥ	‘flea’
ts’iiriiri’	‘hummingbird’
waw’iya	‘run’ (of water)

This Xinkan vowel harmony pattern is interesting typologically for a number of reasons. Vowel harmony systems based on height of mid and high vowels (e.g. an /i, u/ set vs. an /e, o/ set) are very rare, though not unique. Of significance to general claims, however, is the fact that Xinkan harmonic sets of vowels are split based on vowel height, but with /a/ as neutral, occurring with both high vowels and mid vowels. This runs counter to the claim that vowel harmony is due to specific articulatory motivations – it is not possible to talk of a high vowel vs. mid vowel articulatory motivation when the low vowel /a/ occurs with both sets. In addition, the fact that the high central vowel /i/ patterns differently from the other high vowels, and indeed from all other harmonizing vowels, does appear to be unique, and is not known from any other language which has vowel harmony. This has implications for theories of distinctive features, since prevailing views of distinctive features provide no adequate way to showing a natural class of /i, u/ (and /e, o/), while excluding /i/. That is, [+high] does not exclude /i/; [+high, αback, around] is scarcely a satisfying natural class, since it must name three of the four major vowel features (and [-low] need not be mentioned only because of the universal convention that [+high] vowels cannot be [+low], rather than indicating anything revealing about this ‘natural class’). If distinctive features recognized front, central, and back vowels instead of only [-back] (/i/) and [+back] (/i/ and /u/), the problem of excluding /i/ from the /i, u/ harmonic set would not be so difficult. (See Ladefoged and Maddieson (1996: 291–2)

for arguments based on data from other languages.) Knowledge of this unique vowel harmony pattern, with its implications for distinctive feature theory, would never have come to light if all the Xinkan languages had become extinct without documentation.

6.2.1.5 Saami phonemic contrasts

The Saami languages (Uralic family) are spoken in northern parts of Finland, Norway, Russia, and Sweden. Eleven Saami languages are known; two (Kemi and Akkala Saami) are now extinct, and nine are at varying levels of endangerment, with numbers of speakers ranging from 1 to 25,000. Three Saami languages (Ume, Pite and Ter Saami) have fewer than 20 speakers each. The Saami languages have several unusual traits; for example, three different length contrasts for consonants and vowels (Aikio 2008).

The Saami examples presented here are from Aikio (2008), given in the accepted Saami orthography. The three degrees of contrast in consonant length are seen in examples (4–6) from North Saami (western Finnmark dialect):

- | | | | | |
|-----|-----------------|-------|----------|----------------------------|
| (4) | <i>short</i> | [l] | /palū/ | ‘fear’ (genitive) |
| (5) | <i>geminate</i> | [l:] | /pollū/ | ‘wooden bowl’ (genitive) |
| (6) | <i>long</i> | [l::] | /pol:lū/ | ‘wooden bowl’ (nominative) |

Three degrees of vowel length are seen in examples in (7–9) from North Saami (eastern Finnmark dialect):

- | | | | | |
|-----|-----------------|-------|----------------|---------------------------|
| (7) | <i>short</i> | [æ] | /mähte/ | (man’s name) (nominative) |
| (8) | <i>geminate</i> | [æ:] | /mä ähte/ | (man’s name) (genitive) |
| (9) | <i>long</i> | [æ::] | /(in) mää hte/ | ‘(I don’t) know how to’ |

A two-way contrast between short and long vowels is common in languages of the world, but a contrast of three degrees of length is extremely rare and was unknown until fairly recently. Three degrees of vowel length have been reported also for Mixe (Mixe-Zoquean, Mexico) and Yavapai (Yuman, Arizona), and a triple length contrast for both consonants and vowels occurs in Estonian, though conditioned by aspects of morphological structure (Ladefoged and Maddieson 1996: 320). The Saami case is the clearest and puts to rest the once common belief that languages could not contrast more than two degrees of length phonemically (for example, as encoded in the binary distinctive feature [±long] (or in some accounts of vowel length as [±tense] or [±ATR]).

6.2.2 Morphosyntactic structures

6.2.2.1 Word order

A telling example of the value of endangered-language documentation and description is the discovery of languages with OVS

(Object-Verb-Subject) and OSV (Object-Subject-Verb) BASIC WORD ORDER.⁹ In his ground-breaking work on universals, Greenberg (1966[1963]) found only SVO, SOV and VSO basic word orders in the languages of his sample. His Universal 1 reflects this: 'In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object' (Greenberg 1966[1963]: 177). Another version was stated as 'whenever the object precedes the verb the subject does likewise' (Greenberg 1978: 2, Derbyshire and Pullum (1986: 16–17)). Greenberg's sample contained no languages with OVS or OSV basic word order, and this universal as formulated suggests they cannot occur as basic word orders in human languages. However, these orders have now been found as the basic word orders in a few languages, first identified in languages of the Amazon (Derbyshire 1979).¹⁰ An example that became well known is Hixkaryana (Cariban), with only 350 speakers, with OVS order as illustrated in:

- (10) *toto yonoye Kamura*
 man ate jaguar
 'The jaguar ate the man.'

The discovery of these basic word orders as these languages came to be described forced the abandonment of the postulated universal, illustrating forcefully the value of describing and analysing little-known, endangered languages. It is all too plausible that the few languages which have these unusual basic word orders could have become extinct before they were documented, given, for example, the treatment of indigenous peoples of Brazil until recently (and still continuing at the hands of unscrupulous miners, ranchers and logging companies). The *World Atlas of Language Structures* (Haspelmath *et al.* 2005: 330–3) now reports nine cases of languages with OVS basic word order, including Asurini (Tupian; Brazil), Cubeo (Tukanoan; Brazil, Venezuela, Colombia), Hixkaryana (Cariban; Brazil), Mangarrayi (Australian), Pãri (Nilotic; Sudan), Selknam (Chonan; Argentina), Tiriyo (Cariban; Suriname, Brazil), Ngarinjin (Australian), and Urarina (isolate; Peru), and four instances of OSV order, found in Warao (isolate; Venezuela), Nadëb (Makúan; Brazil), Wik Ngathan (Pama-Nyungan; Australia), and Tobati (Austronesian; West Papua, Indonesia). Had all OVS and OSV languages become extinct without being documented or described, linguists would have forever believed in the postulated but erroneous universal about subject preceding object, and on its basis would have made hypotheses about Universal Grammar and about the potentials and limitations of human cognition. This possible but all too plausible case of potential loss of important linguistic information through language extinction illustrates the importance of documenting and describing endangered languages and shows the kinds of contributions they can make to the development of linguistic theory.

6.2.2.2 Nivacé genitive classifiers

Nivacé, mentioned above, has two GENITIVE CLASSIFIERS (also referred to as ‘possessive classifiers’) but no other classifiers. An expression such as ‘my cow’, for example, must use a classifier construction with the ‘possessive domestic animal classifier’. Some examples are:¹¹

- (11) *y-ikla* *waka*
 1SG.POSS-DOMESTIC.ANIMAL.CL cow
 ‘my cow’
- (12) *l-ikla* *kuwayu*
 3SG.POSS-DOMESTIC.ANIMAL.CL horse
 ‘his horse’

Nivacé also has a second possessive animal classifier, *-axe?* ‘prey classifier’ [hunted animal], for example:

- (13) *y-axe* *tašinša*
 1SG.POSS-GAME.CL brocket.deer
 ‘my grey brocket deer (*Mazama gouazoubira*)’

We can contrast (13) with the similar form in (14), where *tašinštax* ‘goat’ (a domestic animal requiring the *-ikla?* genitive classifier when possessed) is derived from *tašinša* ‘grey brocket deer’ by the suffix *-tax* ‘similar to’, whereas *tašinša* ‘grey brocket deer’, because it is ‘game’ requires the *-axe?* classifier for game when possessed:

- (14) *y-ikla* *tašinštax*
 1SG.POSS-DOMESTIC.ANIMAL.CL goat
 ‘my goat’

Genitive classifiers are rare in the languages of the world, occurring primarily in a few South American languages (for example Nadëb of the Makúan family, some languages from the Cariban, Tupí-Guaraní, Jê, Guaicuruan families; Aikhenvald 2000: 147). However, in these languages, the genitive classifier is typically one in a system of noun classifiers, with several other classifiers found in the language. Nivacé (together with some neighbouring languages of the Chaco area) is unusual in that it has no other classifiers, only the unusual genitive classifiers. Such unusual systems need to be studied in more detail to understand their role in classifier systems and in language typology generally. This information could be lost if these endangered languages are not documented; already younger speakers of Nivacé do not know the ‘game genitive classifier’ (Campbell, in preparation).

6.2.2.3 Nivacé demonstratives, tense and evidentiality

Nivacé has an interesting deictic system shown in Table 6.2. (Similar systems are found in some other languages of South America’s Gran Chaco region, also, though most are not well documented.)

Table 6.2. *Nivacle demonstratives*

	Visible	Not visible (known from first-hand experience)	Not visible (hearsay, indirect knowledge)	Moving, deceased
Singular Masculine	na	xa	pa	ka
Plural Human	napi	xapi	papi	kapi
Plural Non-Human	nawa	xawa	pawa	kawa
Singular Feminine	la	la	pa	ka

Though the deictic system itself is interesting, how this system interacts with tense-aspect and evidentiality is of special interest, since it is highly unusual. In spite of having verbs with long strings of affixes, Nivacle verbs carry no grammatical markers of tense or aspect. Rather, tense and aspect are inferred from the demonstratives. For example:

- (15) yoy **na** siwanak
 escape DEM dorado.fish
 ‘the dorado-fish is escaping’ (visible)

- (16) yoy **xa** siwanak
 escape DEM dorado.fish
 ‘the dorado-fish escaped’ (not visible but known from personal experience)

Here, (15) and (16) are identical except for the demonstratives. There is no tense in the verb, but nevertheless tense is inferred from the prenominal demonstratives. In (15), **na** [visible] implies ‘present’; in (16), **xa** [not visible but known] implies ‘past’ (seen previously but no longer present). Such a situation with tense signalled in the noun phrase is not unique (e.g. Nordlinger and Sadler 2004; Tonhauser 2007). Demonstratives in Movima (language isolate, Haude 2006) and Wichí (Matacoan family, Terraza 2008) also signal tense and are the only markers of tense in these languages; however, in these cases the demonstratives bear tense affixes and so the tense is not inferred from the semantics of the demonstratives as it is in Nivacle, but is signalled directly by these affixes. Systems such as this are extremely rare and need to be studied more intensively.

The Nivacle deictic system also interacts with evidentiality. Evidentiality markers indicate the source of knowledge for an utterance and the level of certainty assigned to it. As with tense, there are no verbal morphemes of evidentiality (present in many South American languages), rather the evidentiality is also inferred from the demonstratives. For example, sentences (17) and (18) are identical except for the demonstratives, and yet they are very different in meaning.

- (17) *Boca* *yuʔ-el* ***pa-pi*** *River*
Boca *play-PL* *DEM-PL.HUM* *River*
 ‘Boca play(ed) River’ [two soccer teams in Argentina] (literally ‘Boca played those River’)

The plural demonstrative *pa-pi* [not visible, not known first-hand] shows that the speaker reports this not from first-hand experience, but rather as something reported (hearsay): ‘they say that Boca play(ed) River, but I don’t know this from personal experience and so I do not affirm whether it is true or not’.

- (18) *Boca* *yuʔ-el* ***na-pi*** *River*
Boca *play-PL* *DEM-PL.HUM* *River*
 ‘Boca is playing River’

The plural demonstrative *na-pi* [visible] shows that the speaker sees the event described and therefore affirms it is true: ‘Boca is (truly) playing River’. The meaning of these demonstratives thus also involves evidentiality, accounting for the difference in evidentiality in the two sentences.

An interesting question arises here for linguistic typology: should Nivacle be classified as a language with evidential markers, though evidentiality is only inferred from the demonstratives? In addition, what about cross-linguistic comparisons? Should we consider that these indicators of evidentiality in Nivacle and the evidential systems marked on the verb in some other South American languages are equivalent and thus comparable in some way (see, for example, Epps 2005)?

6.2.2.4 Syntactic ergativity

In some languages subjects of intransitive verbs (bearing a grammatical function we can label as S), and objects of transitive verbs (functioning as O) are marked the same way, and differently from subjects of transitive verbs (functioning as A). This is referred to as MORPHOLOGICAL ERGATIVITY because the special A marking is called ‘ergative’ (and the S/O marking is termed ‘absolutive’). Morphological ergativity is not especially uncommon among the world’s languages. Ergativity is most often reflected only in morphology without major correlates in the syntactic organization of clauses. However, SYNTACTIC ERGATIVITY does exist, and was first reported for Dyirbal (Pama-Nyungan family; Australia), which has only a handful of speakers today (Dixon 1972). Syntactic ergativity refers to syntactic operations which treat the object of a transitive verb (O) and subject of an intransitive verb (S) as a single syntactic category (S/O). It is also sometimes called interclausal ergativity, since it is typically seen in how arguments are treated syntactically in two (or more) linked clauses, for example in relative clause constructions, subordination, and clausal coordination. In Dyirbal, the omission of coreferential

noun phrases in relative clauses or coordinated clauses exhibits syntactic ergativity. In such clauses, the grammatical subject (S) of an intransitive verb can be elided only if it is coreferential with an intransitive subject S or a transitive object O in the (preceding) main clause, as in the following example of coordination:

- (19) *nguma* *yabu-nggu* *bura-n* *banaga-nyu*
 father-ABS mother-ERG see-NONFUT return-NONFUT
 ‘Mother saw father and [father] returned.’ (Dixon 1994: 162.)

In this sentence, with the literal gloss ‘mother saw father and returned’, the second clause is understood to mean ‘father returned’. This is because omission of the subject of an intransitive verb like *banaga-nyu* ‘returned’ requires that it be understood as coreferential with the object of the preceding transitive verb *bura-n* ‘saw’, i.e. *nguma* ‘father’. It is not possible to understand the second clause in (19) with the meaning ‘mother returned’ (notice that in English omission in linked clauses requires coreference of intransitive subjects S or transitive subjects A, so in ‘Mother saw father and returned’ it is understood that ‘mother returned’. English treats S/A as a single syntactic category).

Some aspects of syntactic ergativity have been reported in other languages, though it is uncommon. For instance, a few other Pama-Nyungan languages also have some degree of syntactic ergativity, e.g. Yidiny, Kalkatungu, Warrgamay and Bandjalang (Dixon 1994: 178). Of these, Kalkatungu recently became extinct and the others are critically endangered. Linguistic science would not know syntactic ergativity was possible if all these languages had been lost without being described.

In this section, we have given just a few additional examples of structural features found in endangered languages which have a significant impact on understanding what is possible in human languages. These examples illustrate compellingly the value of documenting and describing endangered languages.

6.3 What can happen to the structure of endangered languages?

Though there are various ways in which languages can become extinct, the most typical is through LANGUAGE SHIFT when a language gradually comes to have fewer and fewer speakers who use it in ever fewer domains until finally no one is able to speak it in any context (see Grenoble, Chapter 2). This process is sometimes called LANGUAGE OBSOLESCENCE, and a language which undergoes it is referred to as an OBSOLESCING LANGUAGE. There can be considerable impact on the structure of the endangered language in these situations (see Campbell and Muntzel 1989). This can have important implications for typological

claims and for the study of language change in endangered languages. In this section we address these issues and the kinds of variation and change found in endangered languages. Specifically, we consider the impact that language endangerment can have on the structure of languages and the kinds of changes and structural differences they can exhibit in contrast to fully viable, non-endangered languages. In 6.3.1 we discuss variation and variability and in 6.3.2 we present types of changes found in endangered languages.

6.3.1 Variation and variability

6.3.1.1 Variation in endangered languages

Variation in obsolescing languages need not exhibit the negative or positive sociolinguistic evaluations usually correlated with social variables such as socioeconomic class, sex, ethnicity, etc. so often found in viable languages. That is, variability often does not bear the social meanings in speech communities undergoing severe language obsolescence that it may elsewhere.

Some changes which take place in endangered language situations are ‘normal’ or ‘natural’ changes which can take place in non-endangerment situations as well (see also O’Shannessy, Chapter 5). An example is the merger of uvular and velar consonants in endangered Mam of Tuxtla Chico (and indeed, also in several non-endangered branches of Mayan, as well). Although imperfect learning may be sufficient to explain many of these cases, the absence of the contrast from the dominant language (Spanish in this case) may also contribute to its loss in the endangered language being replaced.

Some changes in obsolescing languages are natural, but the rate of change can be accelerated, with a change occurring much more rapidly than it might in a healthy language situation. For example, in the Algonkian language Walpole Island Ottawa, the person prefix system exhibits variability and loss which was not known in the language as recently as twenty years ago, and which can be attributed to a natural process of vowel syncope (Fox 2005: 57).

6.3.1.2 Individual variability, the effect of semi-speakers

Most endangered language situations involve gradual decline in speaker numbers and speaker fluency. As more of the community shifts to the dominant language (cf. O’Shannessy, Chapter 5), fewer children learn the minority language, and often those who do so learn it imperfectly, resulting in SEMI-SPEAKERS – people who have learned the language to some degree and are not fully fluent (see Grenoble, Chapter 2). For example, Schmidt (1985a: 381) found that among Jambun Dyirbal (Australia) speakers, variability could be described on a continuum according to the degree of simplification of traditional Dyirbal, and this continuum

correlated with the age of the speakers. This is typical of a gradual shift to a majority language: 'each individual had his own grammatical system for Dyirbal communication, involving simplification of the traditional grammatical norm to a greater or lesser degree.'

Languages can vary greatly in language endangerment situations, with potentially more kinds and greater frequency of variation than encountered in non-endangered languages. Things that are obligatory in a fully viable language may become optional or fail to apply and be lost in the language of semi-speakers. For example, semi-speakers of Tlahuica (a.k.a. Ocuilteco, Otomanguan family, Mexico) sometimes fail to voice stops after nasals (*nt* does not automatically become *nd*, a change which is obligatory in viable Ocuilteco), producing free variation. In Pipil (a.k.a. Nahuate or Nawat, a Uto-Aztecan language of El Salvador), /l/ is always voiceless in final position; in moribund Cuisnahuat Pipil, however, voiced 'l' was allowed also word-finally, with the result of seeming free variation between voiced 'l' and voiceless 'l' in this position. As Swadesh (1934, 1946) observed in his work with the last two speakers of Chitimacha (isolate, Louisiana), glottalized consonants could vary rather freely with their unglottalized counterparts (though original plain consonants did not vary with glottalized ones).

6.3.2 Types of change in endangered languages

6.3.2.1 Normal change (change typically encountered also in non-endangered languages)

As mentioned above, obsolescing languages can undergo ordinary changes that can also be observed to place in languages which are not endangered. For example, in moribund Chiltiupan Pipil, *ts* changed to *s*. The change of an affricate to a fricative (*ts* > *s*) is not an uncommon sound change, and is found in the history of languages that are not endangered. In another example, Pipil speakers today, none of whom are fully fluent, have lost the original vowel length contrast and all long vowels have become short (*V:* > *V*); many have also lost the rule that devoices final *l* (for example /mi:l/ [mi:t] > /mil/ [mil] 'cornfield').

Many changes of this sort may be attributable, at least in part, to influence from the locally dominant language. For example, the change of *ts* to *s* in Chiltiupan Pipil might also reflect influence from dominant Spanish, which has no segment *ts*; the loss of the vowel length contrast and of the voiceless 'l' in final position in varieties of Pipil, though both natural changes, may also reflect the absence of these sounds in Spanish. While influence from the dominant language (or languages) must always be taken seriously into consideration as possibly affecting the structure of endangered languages, we do not emphasize this in this section, concentrating rather on the kinds of structural changes languages can undergo regardless of whether these are abetted by influences from the

dominant language (though see 6.3.2.6. below for discussion of ‘acts of reception’).

6.3.2.2 Overgeneralization of unmarked features (loss of marked features through replacement with unmarked counterparts)

MARKED FEATURES are traits of language which tend to be more unusual cross-linguistically, more difficult for children to learn, and more easily lost in language change. They tend to be replaced by less marked ones (more common cross-linguistically, more easily learned) in language change. That is, difficult contrasts may not be learned, or not learned well. For example, in endangered Mam of Tuxtla Chico (Mayan; Mexico), marked uvular stops were replaced by unmarked velars ($q > k$; $q' > k'$). Chipewyan (Athabaskan; Canada) semi-speakers change glottalized consonants to their plain counterparts ($C' > C$) (Cook 1989). These sorts of changes are also normal and can be found in the history of languages which are not threatened.

6.3.2.3 Overgeneralization of marked features.

In several situations, things that seem ‘exotic’ from the point of view of speakers of the dominant language can come to be overused in unexpected contexts in an obsolescing language. For example, one Jumaytepeque Xinka (Xinkan family, Guatemala) semi-speaker pronounced nearly every consonant as glottalized ($C > C'$). This is not a natural change and would definitely not be expected to occur in fully viable languages. In moribund Teotepeque Pipil, some speakers over-generalized voiceless ‘l’, employing it everywhere at the expense of voiced ‘l’, though in fully viable Pipil, the voiceless ‘l’ is only an allophone of /l/ in final position. In instances such as this, it seems that the semi-speakers are aware of the unusual traits but have not learned where they correctly belong and so use them excessively but inappropriately, as a consequence of imperfect learning.

6.3.2.4 Loss or reduction in phonological contrasts (mergers)

Some instances of phonological merger have already been seen, as for example, Pipil $ts > s$, $V: > V$, (that is, $ts, s > s$; $V:, V > V$), and Mam of Tuxtla Chico $q > k$; $q' > k'$ ($q, k > k$; $q', k > k$). Documentation of both Chitimacha (isolate; Louisiana) (Swadesh 1934, 1946) and Tonkawa (isolate; Texas) (Hojer 1933, 1946) (both now extinct) revealed that the last speakers often merged glottalized consonants with the non-glottalized counterparts.

6.3.2.5 Both overgeneralization and under-generalization

In some instances both overgeneralization and under-generalization can affect structural properties. For example, viable Pipil devices

non-nasal sonorants (*l*, *w*, *y*) word-finally; moribund Teotepeque Pipil, however, overgeneralized voiceless *l*, devoicing *l*'s in all environments, not just final ones, but undergeneralized in the case of *w* and *y* by not devoicing them finally (or anywhere else, for that matter). Through overgeneralization (of voiceless *l*) and under-generalization (of voiceless *w* and *y*) the sonorant final-devoicing process was eliminated. In some other Pipil dialects, as mentioned above, the final *l* also ceased to be devoiced, along with *w* and *y*, meaning that the rule of final devoicing of sonorants was also completely lost in these dialects, though in this case through under-generalization.

6.3.2.6 Acts of reception

ACTS OF RECEPTION in this context refer to instances in which the minority language 'receives', or takes on, traits from the dominant language judged by speakers of the dominant language to be highly valued and also to the avoidance of native traits of the minority language which might be associated with traits of the dominant language which are judged undesirable. Some structural changes can be due to influence from the dominant language in which the minority language takes on highly valued structural traits of the dominant language which are otherwise quite foreign to the minority language. For example, Teotepeque Pipil underwent the change $\xi > r$ (where ξ is a retroflex non-apical laminal fricative, equivalent to a $[j]$ that is retracted to the hard palate). This change is due to an act of reception in which attitudes about pronunciation of the local Spanish of the region are transferred to traits of the minority language, leading it to change. Spanish is the dominant language here, and where as in standard Spanish the phoneme $/r/$ is a voiced alveolar trill, local Spanish has $[\xi]$ as an additional non-standard variant of this phoneme. This variant in Spanish is stigmatized and considered undesirable by local Spanish speakers. This negative evaluation of the $[\xi]$ variant in Spanish was taken over by Pipil speakers and the associated attitude caused moribund Teotepeque Pipil to shift the pronunciation of its native phoneme in order to match the $[\text{r}]$ (trilled 'r') variant in standard Spanish of El Salvador, the prestigious variety. A change of a sibilant such as ξ to a trilled 'r' is highly unusual and unexpected (there are no 'r' sounds in viable Pipil).¹² In another example, some semi-speakers and non-native learners of Pipil often pronounce Pipil initial $/y/$ ($[j]$ in IPA) as $[3]$, as in $[3ek]$ for $/yek/$ 'good', to make it match the $[3]$ prestige pronunciation of initial $/y/$ in Spanish, totally alien to viable Pipil, which has $[j]$ as its only pronunciation. Acts of reception can also influence the lexicon. For example, native words in the endangered language whose meanings are neutral are sometimes avoided or replaced because they sound like words that are obscenities in the dominant language. An example of this is Nivacle *puta* 'rabbit', which sounds like obscene *puta* 'whore' in dominant Spanish. Because this word sounds like an obscenity in the

dominant language, it is replaced in the speech of most Nivacle speakers by *nanxatetax* (derived from *nanxate* ‘hare, jack rabbit’ + *-tax* ‘similar to’). Other examples of lexical avoidance of this sort are not difficult to find.

6.3.2.7 Morphological reduction

Two changes observed with some frequency across obsolescing languages are the decay of case systems and the tendency to change toward more rigid word order.¹³ For example, semi-speakers of American Finnish fail to show adjectives agreeing with nouns in case and number, producing, for example, expressions like *vanha miehe-n* [old man-GENITIVE.SG] ‘the old man’s and *vanha miehe-ltä* [old man-from] ‘from the old man’, where fully competent speakers have *vanha-n miehe-n* [old-GENITIVE.SG man-GENITIVE.SG] and *vanha-lta miehe-ltä* [old-from man-from], respectively, showing agreement. In another example, imperfect speakers of Tlahuica (mentioned above) often eliminate the dual and plural markers which fully fluent speakers do not leave out, as in the following examples where the material missing in the speech of semi-speakers is indicated in parentheses (Campbell and Muntzel 1989: 191–2):

- (20) *kiat-kwe-p-tyii(-nkwe(-βi))*
 FUT-1PL-EXCL-sing(-DUAL(-EXCL))
 ‘We (two, but not you) will sing’

- (21) *kiat-kwe-p-tyii(-hñə(-βi))*
 FUT-1PL-EXCL-sing(-PL(-EXCL))
 ‘We (all, but not you) will sing’

6.3.2.8 Preference for analytic constructions over synthetic ones

Sometimes obsolescing languages prefer analytic constructions over synthetic ones (see Campbell and Muntzel 1989: 192–4).¹⁴ For example, Pipil used to have a synthetic morphological ‘future’:

- (22) (a) *ni-panu-s* [I-pass-FUT] ‘I will pass’
 (b) *ti-panu-ske-t* [we-pass-FUT-PL] ‘we will pass’.

However, later in its more moribund stage, Pipil lost the synthetic morphological ‘future’ and had in its place the analytic syntactic ‘future’:

- (23) (a) *ni-yu ni-panu*
 I-go I-pass
 ‘I will pass’ (literally ‘I’m going to pass’)
 (b) *ti-yawi-t ti-panu-t*
 we-go-PL we-pass-PL
 ‘we will pass’ (‘we’re going to pass’).

In another example, Scottish Gaelic semi-speakers are reported by Dorian (1981: 15) to replace the synthetic conjugated prepositions

(*riu-m* ‘to-me’, *bhu-atha* ‘from-them’) with analytic constructions of free-standing preposition and pronoun (*ri mis* ‘to me’, *bho aid* ‘from them’).

6.3.2.9 Syntactic reduction

Loss of certain grammatical categories and syntactic options, particularly complex sentence constructions, is common in language obsolescence. (The loss of the morphological ‘future’ in Pipil is an example of this.) It is sometimes thought that complex linguistic structures learned later in childhood may be lost because it is at this age that children in many communities often stop using the endangered language. One example is the reduced use and loss of subordinate clauses in dying languages that can be explained in terms of two tendencies: (1) since speakers of moribund languages produce few complex sentences, a child exposed to such language input would have an inadequate model for acquiring them; (2) certain subordinate clauses tend to be used in higher (more formal) styles, but the strong solidarity function of the dying language in some communities emphasizes ‘lower’ (less formal) styles. This may strip away the complex constructions as the ‘higher’ styles cease to be used. When there are competing structures with the same function (meaning), they may tend to be reduced to a single structure in obsolescing language situations.

6.3.2.10 Stylistic shrinkage

Correlated with reduction in grammar is reduction in speech genres and stylistic alternatives (such as verbal art, oral literature, ritual language, formal registers and figurative language) called STYLISTIC SHRINKAGE. Stylistic shrinkage often begins at the formal end of the stylistic continuum, ‘polystylism’ moving to ‘monostylism’, where finally only casual speech remains (Dorian 1980, Hill 1978). However, shrinkage can also take place from the bottom-up, leaving the obsolescing language used only in formulaic ritual settings. Both situations involve stylistic shrinkage with severe reduction in the stylistic range available to speakers. Stylistic reduction correlates closely with syntactic reduction, since different styles are typically characterized by different syntactic options or different frequencies of usages for certain syntactic constructions over others. One example of this is the loss of the morphological future in Pipil (see 6.3.2.8 above), which was characteristic of elevated styles, leaving only the analytic future, more typical of less prestigious speech. Similarly, there may no longer be any Ocuilteco speakers competent in the formulaic ritual language employed in religious ceremonies and marriage petitions. This kind of speech relies heavily on paired couplets (repetition of the same content utilizing alternative semantic and syntactic forms). The last speaker competent in this style may well have been Martha Muntzel’s ritual-language consultant, who died in the mid 1970s (Campbell and Muntzel 1989).

6.4 Conclusion

In this chapter we have presented a sample of the many unique or unusual structural phenomena which have been uncovered in research on endangered languages and which would have been lost to science if these languages had disappeared before they were documented and described. We presented examples of how these discoveries contribute to understanding more completely the full range of what is possible in human languages, challenging some general claims and helping to strengthen others, thus contributing to the development of linguistic theory. We also considered the kinds of variation and change found in endangered and obsolescing languages. We examined the considerable impact language endangerment can have on the structure of the languages involved. We pointed out implications for typological claims, for example that several putative universals and general claims about language must be modified or abandoned based on new findings encountered in various endangered languages. We also considered the kinds of changes encountered in endangered obsolescing languages and some consequences of these for general views about how languages change. All of these considerations show both jointly and individually the importance of documenting and describing endangered languages, and the contributions that findings from these languages can and do make to linguistic theory.

Notes

- 1 Examples are drawn from the literature and from our own fieldwork and first-hand experience with the languages mentioned. This explains why examples from the indigenous languages of the Americas figure so prominently here.
- 2 The Nivacle findings reported here are from Campbell (in preparation), from research supported by the 2003–6 grant ‘Description of Chorote, Nivacle and Kadiwéu: three of the least known and most endangered languages of the Chaco’ funded by the Endangered Languages Documentation Programme, School of Oriental and African Studies (co-principal investigators Lyle Campbell, Verónica Grondona, and Filomena Sandalo).
- 3 It should be noted that many speech sounds have more than one articulatory gesture; however, in cases where the multiple gestures are not articulated relatively simultaneously, it may not be possible to distinguish them phonetically from sequences of segments found in other languages which lack phonotactic evidence for interpreting them as single segments (Ladefoged and Maddieson 1996: 329). This is not the case with this Nivacle sound. Also, the Nivacle sound is different from a velar lateral, a sound which was also unknown until

relatively recently but now known to occur in Mid-Waghi, Melpa, Kanite, Yagaria (New Guinea), Kotoko (Chadic), and Comox (Salishan), where contact is in the velar region with air escaping around both sides of the contact in the region of the back molars (Ladefoged and Maddieson 1996: 190).

- 4 The glottalized stops and affricates are ejective consonants. In the Americanist phonetic tradition, glottalized consonants are represented by an apostrophe following the plain consonant symbol, thus *p'*, *t'*, *k'* ([p^ʔ], [t^ʔ], [k^ʔ] in the International Phonetic Alphabet).
- 5 In Nivacle the voiceless lateral is an approximant, not a fricative (see e.g. Ladefoged and Maddieson 1996: 198).
- 6 Lindblom and Maddieson (1988) are not specific about whether the inventories they have in mind refer only to phonemes or to all the speech sounds (phonemes and their allophones), and of course a more precise definition of what is intended by 'simpler' and 'complex' would be useful. Nevertheless, whether the intent is to refer only to the phonemic inventory or to the entire set of speech sounds (phones) in a language, instances such as Pirahã have both few phonemes and few phones, yet some of the sounds they do have are complex by any definition; for example, Pirahã's voiceless laminal dental plosive followed by a bilabial trill.
- 7 Indeed, many Amazonian languages have become extinct. If we take Loukotka's (1968) count of 1,491 languages once spoken in South America as relatively accurate and compare this with the roughly 420 languages still spoken, we would have to conclude that some 72% of South American languages have become extinct. Probably Loukotka's list includes some names that do not actually correspond to real languages and in some cases multiple names referring to the same language; however, even allowing for such problems, the number of languages which have become extinct, a great many with no documentation or description, is very large.
- 8 The investigation of Xinkan was supported in part by a grant from the National Science Foundation to the University of Utah, 'Xinkan, Pipil, and Mocho': Bringing Three Endangered Language Documentation Projects to Completion' (co-Principal Investigators: Lyle Campbell, Laura Martin, Terrence Kaufman.)
- 9 The term 'basic order' is identified with the order that occurs in stylistically neutral, independent, indicative clauses with full noun phrases for the S (Subject) and O (Object) arguments in transitive clauses. Basic order and dominant order are usually equated, though not always by all scholars (see Siewierska 1988: 8). To say a language has a basic word order of a particular kind does not mean that other orders are not possible within it.
- 10 VOS order was also not found in Greenberg's original sample of languages and was also excluded by his proposed universal, though a number of examples of VOS languages soon came to be identified.

- 11 Abbreviations used in the glosses are: ABS absolutive; CL classifier; DEM deictic; EXCL exclusive; ERG ergative; FUT future; HUM human; NONFUT non-future; PL plural; POSS possessive; SG singular.
- 12 Rhotacism might be mentioned, but this is not the form that rhotacism typically takes in languages where it is known, usually involving intervocalic /s/ and a non-trilled 'r'.
- 13 Similar changes, of course, can take place in non-obsolescing languages, but it is quite common to see such changes and for them to take place rapidly in obsolescing languages, generally more commonly and rapidly than in non-obsolescing languages. Naturally, the dominant language can influence decay of case and word-order change to reflect the character of the dominant language more, and the presence of case marking in the dominant language could retard decay of case in the obsolescing languages.
- 14 Changes towards analytic structures can also take place in non-obsolescing languages, but not typically as frequently nor as rapidly as in obsolescing languages.